VICTAS
Vitamin C, Thiamine and Steroids in Sepsis

A multi-center, randomized, placebo-controlled, double-blind, adaptive clinical trial of vitamin C, thiamine and steroids as combination therapy in patients with sepsis

Statistical Analysis Plan

Version 2.0
2 August 2019

Christopher J. Lindsell, PhD

2 August 2019
Date

1. INTRODUCTION

This document describes the statistical analysis plan (SAP) for a double-blinded, randomized 1:1 placebo-controlled adaptive sample size clinical trial to investigate the efficacy of the combined use of vitamin C, thiamine, and hydrocortisone for patients with sepsis (VICTAS). Initial enrollment will include up to 500 participants and may continue to a maximum of 2000 participants if no stoppage rules are triggered. Stoppage rules have been defined *a priori* and will determine the final number of participants enrolled in the trial. Participants will be aged 18 years and older with suspected or confirmed sepsis that are either admitted to or awaiting admission to an Intensive Care Unit (ICU). The presence of sepsis will be evidenced by (1) the ordering of blood cultures and the administration of at least one antimicrobial agent, and (2) acute respiratory and/or cardiovascular organ dysfunction that is attributed to the sepsis event Randomization to either the study invention or control must occur within 24 hours of onset of the qualifying organ dysfunction, and treatment must be started within 4 hours of randomization. As many as up to 70 sites may participate in subject recruitment and enrollment.

This SAP describes the approach to analyzing primary, secondary, and exploratory endpoints, as well as safety endpoints, once the trial is stopped. Details of the design and adaptations, including sample size justification, interim analysis plans, and the approach to multiplicity during interim looks at the data has been previously described.[1]

2. TREATMENT ARMS

Participants will be randomized 1:1 to receive either intervention or control, defined as follows:

a) **Intervention**: Intravenous vitamin C (1.5 grams), thiamine hydrochloride (100 mg), and hydrocortisone sodium succinate (50 mg) administered within four hours of randomization and then every 6 hours thereafter. The intervention continues until either ICU discharge or 16 drug administrations have been completed (96 hours), whichever is first. On days a patient is treated with open label hydrocortisone at a dose of ≥200 mg/day (or equivalent), study drug 3 (hydrocortisone succinate) is withheld. If a subsequent daily medication check during the treatment period reveals that open label hydrocortisone has been decreased to < 200 mg/day (or equivalent), study drug 3 will be administered.

b) **Control**: Matching placebos are administered intravenously within four hours of randomization and then every 6 hours thereafter. Placebo administration continues until either ICU discharge or 16 drug administrations have been completed (96 hours), whichever is first. On days a patient is treated with open label hydrocortisone at a dose of ≥200 mg per day (or equivalent), study drug 3 (placebo) is withheld. If a subsequent daily medication check during the treatment period reveals that open label hydrocortisone has been decreased to < 200 mg/day (or equivalent), study drug 3 will be administered.

3. ENDPOINTS

Primary endpoint

The primary endpoint for this trial is ventilator and vasopressor-free days (VVFD) on day 30 following midnight of the randomization day. This is computed as a backwards count of consecutive whole days free of both respiratory and vasopressor support between day 30 and the most recent use of either respiratory or vasopressor support (Figure 1). Note that days free of respiratory and vasopressor support that occur between periods with support do not count towards VVFD. The day of randomization is day 0 and the next calendar day is day 1. Day 0 will not contribute to the count because participants must require respiratory support or vasopressors to be enrolled. For a day to

count as free of both respiratory and vasopressor support, a patient cannot receive any of the following on that calendar day:

- Mechanical ventilation via an endotracheal tube or tracheostomy tube
- Non-invasive positive pressure ventilation with supplemental oxygen
- High flow nasal cannula at ≥40 l/min with an FiO2 ≥0.4
- Any vasopressor

Participants in need of respiratory or vasopressors support on day 30 will be assigned 0 VVFDs. Participants who die before day 30 will also be assigned zero VVFDs. For participants alive but not observed to day 30 (e.g. participants who are discharged or transferred to another facility), the last observed status will be carried forward. That is, if the participant was last seen on respiratory or vasopressor support, it is assumed they remained that way until day 30 and will thus be scored zero VVFDs. If the participant was last seen not requiring respiratory or vasopressor support, it is assumed they stayed that way until day 30 and the unobserved days will count as VVFDs.

We note that it is possible a patient who is discharged might experience a subsequent hospitalization during which respiratory or vasopressor support is provided. However, we are unlikely to have knowledge of such events and they will not be counted against the VVFD for the original enrollment; status at discharge for the first hospitalization will be carried forward.

The VVFD is an ordinal variable, with death or requirements for vasopressor use or respiratory support resulting in lower scores, and sustained absence of vasopressors and respiratory support resulting in higher scores.

Key secondary endpoint

30-day mortality

The secondary endpoint for this trial is death within 30 days. The 30 days begins at midnight on the day of randomization (day 0). A death within 30 days will count towards this endpoint. Death is a binary variable. As with VVFD, we will use the last observed value carried forward in the case a patient is discharged prior to 30 days. That is, a patient discharged alive will be assumed alive at 30 days.

Additional exploratory endpoints

- Change in the SOFA score between baseline and day 4
- A binary variable indicating mortality between the time of randomization and 180 days following midnight following randomization
- A binary variable indicating mortality between the time of randomization and departure from the ICU
- Length of ICU stay, measured in days, from the midnight following randomization to day of departure from the unit; a partial day will count as a whole day
- Length of hospital stay, measured in days, from the midnight following randomization to day of departure from the hospital, a partial day will count as a whole day
- Renal replacement-free days (RRFD) at 30 days, computed as a backwards count of consecutive whole days free of renal replacement therapy beginning at 30 days from the midnight following randomization, i.e. calculation will follow the same rules as VVFD.
- ICU delirium, measured as number of whole days alive and free of both delirium and coma between midnight on the day of randomization and day 5. Delirium and coma are assessed

using the Confusion Assessment Method-ICU (CAM-ICU). All coma and delirium free days count towards this endpoint regardless of whether they are consecutive or not. If multiple assessments are done on a single day, all assessments must be free of delirium and coma for the day to count towards the endpoint.

- Neuro cognitive outcomes at 180 days measured among survivors using the following instruments:
    - Attention (WAIS-IV Digit Span)
    - Delirium (Telephone Confusion Assessment Method)
    - Executive Functioning (Hayling Test)
    - Language (Controlled Oral Word Association Test or COWA)
    - Memory (Paragraph Recall from the WMS- IV)
    - Orientation (Telephone Interview for Cognitive Status)
    - Reasoning (WAIS-IV Similarities)
    - Activities of Daily Living (Katz ADL)
    - Employment (Employment Questionnaire)
    - Instrumental Activities of Daily Living (Functional Activities Questionnaire)
    - Depression (Beck Depression Inventory-II),
    - PTSD (Post-Traumatic Stress Disorder Checklist for the DSM V)
    - EuroQol, 5 dimension (EQ5D)

Safety endpoints

Potentially Associated Adverse Events (PAAEs)

Safety endpoints to be included in the safety analysis for this trial include all potentially associated adverse events (PAAEs). The pre-specified PAAEs are:

- Nephrolithiasis
- Hemolysis
- Hypersensitivity Reactions
- Injection Site Reactions

In addition, we will report other potentially associated adverse events that are not listed above.

Adverse events

Due to the nature and clinical course of patients with sepsis and septic shock, a substantial number of adverse events are expected among participants, including but not limited to:

- Death
- Renal failure
- Respiratory failure
- Heart failure
- Pneumonia or other / new infection
- DVT or PE
- Complications related to ICU procedures
- Arrhythmia
- Delirium
- Bowel ischemia
- Ileus
- Leukopenia or leukocytosis

- Anemia or thrombocytopenia
- Coagulopathy (DIC)
- Hypoglycemia
- Electrolyte abnormalities

These adverse events are common in sepsis and septic shock and are thus not expected to reflect safety of the treatment regimen. Conversely, absence of these events is expected to contribute to efficacy outcomes and several are included as efficacy endpoints. There is no plan to summarize or report these events to characterize safety.

4. DESIGN CONSIDERATIONS

Randomization

Participants are randomized in a 1:1 ratio to receive either intervention or placebo. Randomization will use permuted small blocks of random size, stratified within site. No other stratification or control for imbalance will be used. The randomization schema will be deployed via the central investigational pharmacy.

Adaptations and stopping

Adaptations and stopping rules are described more completely elsewhere.[1] They are briefly reviewed here for context. The trial is designed to detect a moderate effect on the primary endpoint of VVFD, while allowing early stoppage if a very large effect is observed on the secondary endpoint of mortality. As described in the adaptive design, early interim analyses will be performed to compute the predicted probability of success on the mortality endpoint with N=200, 300, and 400 enrolled participants. Interim analyses will include all data, monitored and unmonitored, for completed participants, as well as information on the number of patients enrolled who do not yet have outcomes available. Efficacy stopping rules, but not futility stopping rules, are in place for these early interim analyses. If a sufficiently large effect is observed, accrual will be stopped, all currently enrolled participants will be followed for outcomes and the primary analysis will focus on the mortality endpoint.

If the trial proceeds beyond N=400, additional interim analyses will be conducted at N=500, 1000, and 1500 based on both VVFD and mortality. Futility and efficacy rules are defined for these interims. Once a stopping rule has been triggered or when 2000 patients have been recruited, accrual will stop, all enrolled participants will be followed for outcomes and the primary analysis will focus on the VVFD endpoint.

Power and Sample Size

Power and sample size considerations are described elsewhere and are included here for completeness1 If the intervention truly causes a mortality difference of 20%, study power is approximately 99%, and the trial would be stopped before 500 participants are enrolled with very high probability (>95%) of success. If the intervention has a more moderate effect on mortality, the trial will most likely continue beyond 500 participants. For a true mortality difference of 5% and true average improvement of 0.6 days free of ventilators or vasopressors for participants that do not die, the power of the study is approximately 95%. The primary outcome assessed when N= 500, 1000, 1500, or 2000 participants is ventilator and vasopressor-free days. The overall Type I error rate for the trial will be controlled at 2.5%. Thus, the early interim looks at N=200, 300, and 400 are designed to conservatively spend alpha so that 2.4% remains for the analysis 500 or more enrolled participants.

5. DEFINITION OF ANALYSIS SETS

Intent-to-treat Analysis Set

All randomized participants will be included in the intent-to-treat analysis set. The intent-to-treat participants will be used for all primary, secondary, and other efficacy analyses. In these analyses, participants will be classified according to the treatment to which they were randomized, regardless of what treatments or how many study treatments were given.

Participants who withdraw consent will be included in the intent-to-treat analysis set. If, at the time consent was withdrawn, the participant gave consent for observation of outcomes then observed outcomes will be used. Otherwise, the last observed value will be carried forward.

Per Protocol Analysis Set

All participants who are included in the intent-to-treat analysis set who correctly receive at least four doses of assigned study treatment (all three components of the study drug or placebo, adjusted for open-label steroids) and did not incur any major protocol deviations or violations will be included in the per protocol analysis set. In this analysis, participants will be classified according to the treatment they received. Major protocol deviations or violations will be identified prior to the unblinding of the study at the final analysis and will include:

i) Found to violate any inclusion or exclusion criterion
ii) Condition adjudicated not to be sepsis
iii) Received one or more doses of the unassigned study treatment
iv) Study hydrocortisone (or placebo) not adjusted for use of open label steroids.
v) Other protocol deviations classified as ‘major’ by majority vote of the Executive Committee, who shall be blinded at the time of the vote.

Safety analysis set

Participants who are randomized and receive at least one administration of study treatment will be included in the safety analysis set. If a subject received both placebo and active treatment, they will be considered as having received active treatment. All other participants will be classified as not having received active treatment.

6. ANALYSIS

Timing of Analysis

Once a decision to stop the trial has been made, the primary analysis may proceed after all enrolled participants have completed the 30-day follow up, 30-day data have been monitored, and 30-day data are declared query free. All data up to and including day 30 will be locked at this time. Analysis of additional efficacy endpoints and long-term outcomes will proceed after all enrolled participants have completed 180 day follow up, the 180-day data have been monitored, the 180-day data are declared query free, and the remainder of the database has been locked.

Blinding

Trial investigators and research teams are blinded to treatment assignment. There are two groups of study statisticians, one of which is performing the interim analyses and one of which is conducting the primary study analyses. Neither group of statisticians will be blinded, and both groups may be in

attendance during closed DSMB deliberations. This statistical analysis plan was drafted prior to the first interim analysis and prior to unblinding.

Descriptive analysis

Using data pooled across all sites, the study sample will be characterized based on demographic and clinical variables measured at randomization, unless otherwise indicated. Specifically, the following variables will be described:

- Age (years)
- Race (African America, Caucasian, Other)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino, or not reported)
- Sex (Male or female)
- Education (less than high school, high school or GED, some college)
- BMI (kg/m^2)
- Medical history (yes, no)
  - diabetes
  - cardiovascular
  - neurological
  - respiratory illness
  - current cancer
- Eligibility criterion (ventilator, vasopressor, both)
- Source of admission (ED, intermediate care or step down unit, floor, other)
- Admission reason (Sepsis, other medical, urgent surgical [necrotizing soft tissue, bowel obstruction, bowel ischemia, burn, trauma], other surgical)
- Baseline Vitals (closest measurement prior to time of randomization)
  - Heart rate (BPM)
  - SBP (mmHg)
  - DBP (mmHg)
  - Mean arterial pressure (mmHg)
  - Respiration rate
  - Temperature (°C)
- Baseline labs (closest measurement prior to time of randomization)
  - White blood cell count (K/cu mm)
  - Platelets (K/cu mm)
  - Hemoglobin (g/dL)
  - Lactate (mmol/L)
  - Creatinine (mg/dL)
- Baseline severity
  - APACHE II (continuous score)
  - SOFA (continuous score)
  - CAM-ICU (delirium present or absent))
- Infection

- Infection source (Lung, blood or vascular access, urinary tract, intra-abdominal, skin or soft tissue, CNS, bone or joint, other, unknown*; use latest confirmed source (discharge or day 30 > ICU discharge > baseline). If no confirmed source available, use last presumed source*)
- Gram positive organism; *use latest (discharge or day 30 > ICU discharge > baseline).*
- Gram negative organism; *use latest (discharge or day 30 > ICU discharge > baseline).*
- Fungal infection; *use latest (discharge or day 30 > ICU discharge > baseline).*
- Organism not identified; *use latest (discharge or day 30 > ICU discharge > baseline).*
- Other infection; *use latest (discharge or day 30 > ICU discharge > baseline).*
- Unknown infection; *use latest (discharge or day 30 > ICU discharge > baseline).*

Categorical variables will be described using frequencies and proportions. Continuous variables will be described using mean and standard deviation, as well as using median and interquartile ranges (IQR). The sample will be described overall and stratified by group assignment according to the intent-to-treat principle. No statistical testing will be done to compare characteristics between groups.

Main Analysis

The main analysis will be a simple comparison between the two treatment groups according to the intent-to-treat principle. If the study stops before N=500, the first analysis will be based on mortality. Otherwise, the first analysis will be based on VVFDs.

VVFDs: A Wilcoxon rank-sum test (i.e. Mann-Whitney U test) will be used to compare VVFDs between treatment groups using a one-sided alpha of 0.022. As described in the adaptive design report, this threshold controls the type 1 error accounting for multiple analyses at N=500, 1000, 1500 and 2000.[1] If the sample size is N<500, the VVFD endpoint will be tested only if the mortality endpoint is successful.

Mortality: If the study stops prior to N=500, the mortality endpoint will be tested first with a chi-square test using a one-sided alpha=0.001 (i.e. 0.1%). If the study reaches N=500 or more, mortality will be compared between treatment groups only if there is a difference observed on VVFDs. In this case, a one-sided alpha of 0.024 will be used.

Description of endpoints: Endpoints will be described using median and interquartile range for VVFDs and using frequency and percent for mortality. The distribution of VVFDs will be described using histograms. Mortality point estimates will be reported with 95% confidence intervals. Descriptions will be given overall, and for each treatment group. Differences in median VVFDs will be computed, with 95% confidence intervals for the difference. Similarly, differences in proportions for mortality will be calculated with 95% confidence intervals. Confidence intervals will be calculated using bootstrapping.

Sensitivity analysis

Our sensitivity analyses are not designed to preserve type I error rates. They are designed to explore possible sources of bias that might inform interpretation of the main analysis. As such, all sensitivity analyses we will use a two-sided alpha of 0.05. We will also emphasize the magnitude and confidence intervals of differences over statistical significance.

Per protocol analysis: We will duplicate our main analysis using the per protocol dataset.

Missingness: Due to the method of using last value carried forward to assign unobserved outcomes, there will be no missingness on the primary outcomes for the main analysis. We will conduct one sensitivity analysis in which we will replicate the main analysis but include only those participants with observed outcomes.

***Steroids*:** Since participants in either arm can receive open label steroids, we will replicate the primary analysis excluding those who were treated with open label steroids in the placebo group.

Safety Analysis

This study is not designed to test safety. No statistical comparison of safety will be done. Safety endpoints will be reported in tabular format, grouped by whether the participant received any active treatment or not.

Adjusted analysis

We will use two approaches to estimate treatment effects adjusted for covariates. Generalized linear mixed models will be used to estimate the conditional effect of treatment with site as a random effect. Generalized estimating equations will be used to estimate the marginal effect of study treatment. Mortality will be modeled assuming a logit link function. A proportional odds model will be specified for VVFDs. Models will consider baseline variables as listed in the descriptive analysis. Multiple imputation based on predictive mean matching will be used to overcome any missingness in covariates. Restricted cubic splines will be used for addressing potential non-linearities in the association between continuous variables and outcomes. Interaction terms will not be considered in the main adjusted analysis. Best fit models will be constructed based on the intent-to-treat principle. Effect sizes and overall fit will be emphasized for selecting the final model; statistical significance alone will not be used to decide which variables are included in the best fit models. The best fit models will then be applied to the per protocol analysis set in a sensitivity analysis. In addition, if any multiple imputation was required, the best fit models will be applied to complete cases.

Differential treatment effects and subgroup analysis

Using the best fit models from the principal adjusted analysis, we will evaluate the interaction between treatment group assignment and each of the following variables: sex, race, ethnicity, age, admission reason, infection, baseline illness severity, baseline lactic acid, and eligibility criterion. Interactions will not all be tested at once within the same model but will be tested one by one by adding the interaction term to the best fit model. If any interaction achieves a p-value of ≤0.2, we will report treatment effects within each subgroup defined by that interaction variable.

Analysis of exploratory endpoints

A number of exploratory endpoints have been specified. We will proceed with exploring the effect of treatment on these endpoints in a similar manner as for the primary and secondary endpoints:

i) Endpoints will be described as medians and interquartile ranges or frequencies and proportions; histograms will be generated for ordinal variables

ii) Binary variables will be compared using a chi-square test and ordinal variables will be compared using a Wilcoxon rank-sum test.

iii) Differences and 95% confidence intervals of differences between study arms will be computed
iv) Exploratory endpoints will be modelled with adjustment for baseline covariates
v) Differential treatment effects will be evaluated and consequent subgroup effects reported

All analyses of exploratory endpoints will proceed under the intent-to-treat principle. A p-value of 0.05 will be used, although emphasis will be placed on effect sizes. We do not expect exploratory endpoints to be continuous and normally distributed. However, any outcome variable that meets these criteria may be compared between treatment groups using student's t-test and a linear link function will be used for modeling purposes. It is possible that exploratory endpoints may be missing. Analyses of exploratory endpoints in the presence of missing outcomes will proceed using complete case analysis. A sensitivity analysis using multiple imputation for missing outcomes will be conducted

7. SUMMARY

The analyses described here are those necessary **to answer the trial's primary** question of whether combined treatment with vitamin C, thiamine and steroids is more effective than placebo in increasing days alive and free from respiratory and vasopressor support and reducing mortality in patients with sepsis.

Beyond our analysis exploring the effect of treatment on primary, secondary, and exploratory endpoints, we expect there to be multiple additional exploratory analyses conducted. It is not possible to predetermine the nature of such analyses, particularly as a rich biospecimen repository is being developed as a component of this study. However, we are committed to preserving rigor and reproducibility and will pre-specify each subsequent analysis in the context of the specific question to be answered, cognizant of bias and missingness in the data.

References

1. Hager DN, Hooper MH, Bernard GR, Busse LW, Ely EW, Fowler AB, Gaieski DF, Hall A, Hinson JS, Jackson JC, Kelen GD, Levine M, Lindsell CJ, Malone RE, McGlothlin A, Rothman RE, Viele K, Wright DW, Sevransky JE, Martin GS. The Vitamin C, Thiamine, and Steroids in patients with Sepsis (VICTAS) Protocol: A prospective, multi-center, double-blind, adaptive sample size, randomized, placebo-controlled, clinical trial. *TRIALS* 2019;20(1):197